心理科学进展　2020, Vol. 28, No. 2, 252–265
Advances in Psychological Science　DOI: 10.3724/SP.J.1042.2020.00252

基于脑结构像的精神分裂症机器学习分类*

郑　泓[1,2]　蒲城城[3]　王　毅[1,2]　陈楚侨[1,2]

([1] 中国科学院心理研究所心理健康重点实验室，神经心理学与认知神经科学研究室，北京　100101)
([2] 中国科学院大学心理系，北京　100049) ([3] 北京大学第六医院，北京大学精神卫生研究所，
国家卫生健康委员会精神卫生学重点实验室(北京大学)，
国家精神心理疾病临床医学研究中心(北京大学第六医院)，北京　100191)

摘　要　将机器学习应用于精神疾患的临床和基础研究是近年来的趋势。研究者将机器学习应用于精神分裂症患者及高危人群的T1加权像和弥散张量成像的脑影像数据中，为了解疾病的生理病理学机制提供帮助。回顾以往研究发现额叶及颞叶的脑结构特征具有较高的区分能力，行为数据和脑影像数据结合的分类效果优于单模态数据。现阶段研究存在样本量不足和泛化能力欠缺的局限，未来研究应注意扩大样本量、制定标准化的分类方法，从而进一步探究机器学习在精神疾患中的作用。

关键词　脑结构像；弥散张量成像；机器学习；精神分裂症；高危人群
分类号　R395

精神分裂症是一种严重的精神障碍，患病率约为1%，患者表现出思维方式和行为举止怪异、感知觉扭曲、认知失调等症状并且常伴有情感迟钝、社会退缩等特征(Carpenter & Buchanan, 1994)。近年来，研究者利用脑影像技术发现精神分裂症患者的脑结构存在异常，包括灰质体积和密度改变，白质异常和脑室体积增大等，为理解精神疾病的病理机制提供支持(Bakhshi & Chance, 2015; Palaniyappan, 2017)。现阶段精神分裂症脑影像研究多采用临床患者和健康对照的组间均值差异检验来考察患者的异常特征，然而这一方法获得的结果无法应用到个体水平上，机器学习(Machine Learning)技术可以弥补以上不足。机器学习这一概念最早由IBM (International Business Machines Corporation)公司的 Arthur Samuel (1959)提出。Tom Mitchell提出更广泛且正式的机器学习定义："对于某类任务T和性能度量P，如果一个计算机程序在T上以P衡量的性能随着经验E而自我完善，那么称这个计算机程序在从经验 E 学习" (Mitchell, 1997)。在精神疾病领域，研究者利用机器学习从脑影像数据中学得规律、构建模型并利用模型在个体水平上进行预测，同时发现最具有鉴别能力的特征。

目前已经有研究者利用机器学习分类方法，尝试探讨如何从人群中识别出精神分裂症患者(Ardekani et al., 2011; Nieuwenhuis et al., 2012; Winterburn et al., 2017)、从高危人群中识别出未来转变为精神分裂症患者的个体(Zarogianni, Storkey, Johnstone, Owens, & Lawrie, 2017)以及区分处于不同疾病阶段的临床患者(Guo, Palaniyappan, Liddle, & Feng, 2016)。在训练模型的过程中，模型选取的特征一般是能够区分各类别的显著特征，因此精神分裂症患者、高危人群和健康对照的机器学习分类研究可以为精神分裂症的疾病诊断、预后判断、疗效判断的相关生物标记物的寻找提供帮助。

本文将从以下四个方面进行系统论述：首先简要阐述机器学习的主要步骤，其次对以往利用脑结构像对精神分裂症患者及高危人群开展的机器学习研究进行系统梳理，并讨论现阶段精神分裂症领域机器学习研究的局限，最后提出未来研究方向。

收稿日期：2019-01-28
* 国家自然科学基金项目(31871114, 31400884)。
通信作者：王毅, E-mail: wangyi@psych.ac.cn

1　机器学习的主要步骤

将机器学习应用于精神分裂症的脑结构像(structural Magnetic Resonance Imaging, sMRI)数据时，主要步骤包括从脑影像中提取特征数据、特征降维和特征选择、训练模型、评估模型以及利用模型预测新样本(见图 1)。

1.1　提取数据

在精神分裂症的脑结构像研究中，采用高信噪比的大脑图像是分析的基础。研究者应进行图像质量控制，以保证用于机器学习数据的可靠性和有效性。从脑影像中提取数据后，数据集将被分割成训练集和测试集或者训练集、验证集和测试集。研究者使用训练集数据寻找规律、生成模型，利用验证集数据调整模型参数、选择最优模型，最终使用测试集数据检验最终模型的性能。

1.2　特征降维和特征选择

特征降维和特征选择能有效提高模型的泛化能力，是数据处理的重要步骤，对高维数据的处理尤为重要。特征降维是高维数据转变成低维数据的过程，常用方法是主成分分析(Principal Components Analysis, PCA)。特征选择是从原始或降维后的数据中选取具有最佳分类效果的特征子集的过程，该子集的类别分布应与原始数据的类别分布相似。Guyon, Weston, Barnhill 和 Vapnik (2002)基于评价准则将特征选择方法分为过滤法(Filter)、包装法(Wrapper)和嵌入法(Embedded)，三类特征选择方法在精神分裂症的脑影像分类研究中均有应用。过滤法根据特征的统计指标或特征与类别之间的相关性对特征进行评分，根据设定的阈值或设定的个数选择特征，有助于降低特征之间的相关性，提高特征与类别之间的相关性。包装法是根据算法的分类性能增减特征，如使用基于支持向量机的递归特征消除方法(Support Vector Machine-Recursive Feature Elimination, SVM-RFE)选择特征。嵌入法合并特征选择和分类器训练过程，在训练过程中自动选择最优特征子集，如决策树算法在训练过程中优先选择分类能力更强的特征。除此之外，研究者还可以根据以往文献选择或移除特征，例如将精神分裂症患者受药物影响较大的纹状体的影像数据移除(Nieuwenhuis et al., 2012)。

1.3　训练模型

研究者利用机器学习算法寻找脑影像数据规律并生成模型。机器学习算法一般分为监督学习和无监督学习，区别在于算法是否使用样本数据的分类信息或目标值。在精神分裂症的研究中，研究者通常使用支持向量机(Support Vector Machine, SVM)、线性回归(Linear Regression)等监督学习算法探究精神分裂症患者或高危人群和健康对照的分类问题。

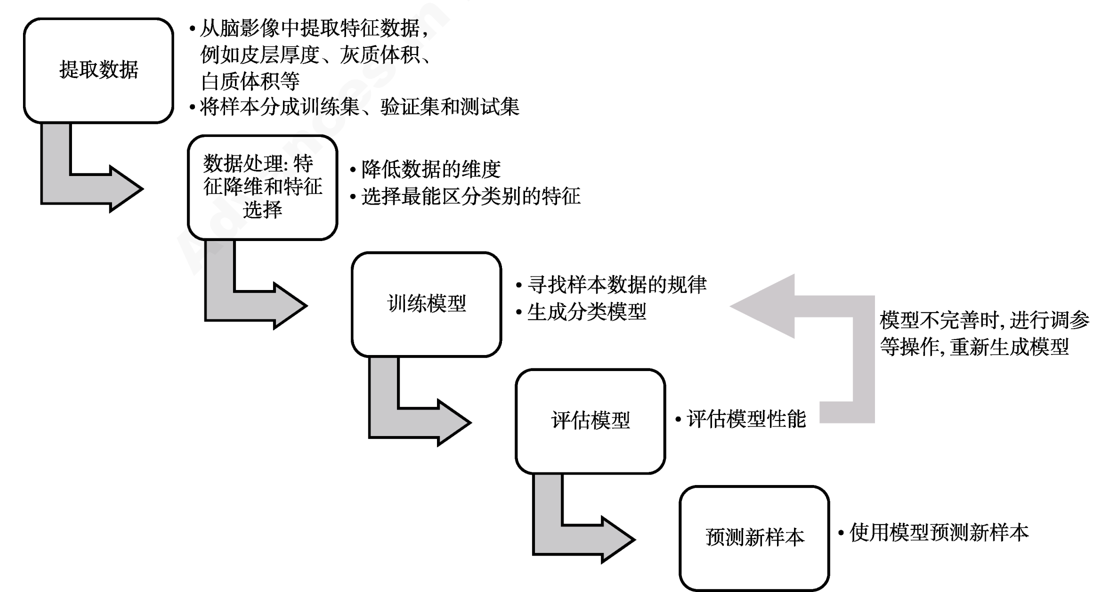


图 1　脑影像数据的机器学习分析步骤

1.4 评估模型和利用模型预测新数据

泛化能力是指模型适用于新数据的能力，研究者利用性能度量来衡量模型泛化能力，评估模型优劣(周志华, 2016)。在精神分裂症的研究中，正确率(Accuracy)、敏感性(Sensitivity)即查全率(Recall)、查准率(Precision)和特异性(Specificity)是最直观的性能度量。正确率是判断正确的样本数占样本总数的百分比。敏感性指所有阳性样本中真阳性样本的比例。查准率指真阳性样本占判定为真的样本的百分比。特异性指真阴性样本占所有阴性样本的百分比。根据性能度量结果，研究者采取更改算法、调节模型参数等方法调整模型。最终模型确立后，可用来预测新数据的类别。

1.5 交叉验证

交叉验证是特征选择、参数调整和最终模型检验中评估模型性能的有效方法，数据处理操作(如特征降维和特征选择)必须嵌入在交叉验证中。K 折交叉验证(K-fold cross validation, K-fold CV)是最常见的交叉验证方法，它将样本分为 K 份，每次取 K−1 份为训练集，剩下 1 份为测试集，将平均正确率作为最终预测结果。在大多数研究中，研究者选择使用 5 折或 10 折交叉验证，也可根据样本大小和数理统计标准确定 K 值(Dwyer, Falkai, & Koutsouleris, 2018)。由于受到样本量的限制，脑影像的机器学习研究通常采用 K 折交叉验证法的特殊形式——留一法(Leave One Out Cross Validation, LOOCV)，其 K 取样本总数 N。留一法虽然简单，但具有高偏差、高耗时等缺点。

另外，有研究者建议使用嵌套交叉验证(Nested Cross Validation) (Madsen, Krohne, Cai, Wang, & Chan, 2018)。嵌套交叉验证包括评估模型性能的外循环和用于选择最优特征、最优参数的内循环，令小样本数据集能够采用不同的数据来评估模型和选择特征、调整参数，最大程度减少偏差，缓解小样本数据的可用性问题，提高评估结果的可信度。目前已有多篇脑影像数据机器学习研究使用嵌套交叉验证方法(Borgwardt et al., 2013; Dyrba, Grothe, Kirste, & Teipel, 2015; Peng et al., 2017; Zarogianni et al., 2017)。

2 基于脑结构特征的精神分裂症患者与健康对照的分类研究

我们对 2010 年以来精神分裂症领域采用脑结构像数据(包括 T1 加权像、弥散张量成像(Diffusion tensor imaging, DTI))的机器学习研究进行文献搜索，使用的数据库包括 Web of Science, PubMed, Elsevier，采用的关键词包括“structural MRI”, “Machine Learning”, “Support Vector Machine”, SVM, schizophrenia, “high risk”, schizotype 以及“voxel based morphometry”, VBM, “grey matter”, “gray matter”, DTI, “diffusion tensor imaging”的不同组合。表 1 中列出截止到 2019 年 1 月 17 日搜索到的 34 篇文献的主要方法和结果。

根据分类特征是否来源于同一模态，脑影像数据分为单模态数据和多模态数据。为探讨上述两种数据和不同脑结构特征对精神分裂症脑影像的机器学习分类的影响，下文将按照基于脑结构像的单模态机器学习研究、基于 DTI 的单模态机器学习研究、多模态机器学习研究、脑影像数据和行为数据相结合的机器学习研究这一顺序进行综述。

2.1 基于脑结构像的单模态机器学习研究

基于脑结构像的机器学习研究中，最常用的特征包括大脑灰质体积、白质体积和皮层厚度。在这些研究中，以慢性精神分裂症患者、首发精神分裂症患者，以及精神病高危个体为对象的分类研究均有报道。从机器学习的分类表现上看，现有研究的分类正确率在 36.84%到 98%之间；在特征方面，使用额叶、颞叶的脑结构特征可以达到较高的正确率。

2.1.1 精神分裂症

在精神分裂症的分类研究中，总样本量最小值为 39，最大值为 606，多数研究总样本量超过 100；分类正确率在 44.74%到 98%之间；研究结果指出，额上回、额中回、额下回、颞中回、颞上回、梭状回、海马、楔前叶、丘脑和枕叶的脑结构特征具有较高的分类正确率；算法对分类正确率的影响不明显，特征选择方法可能会影响分类表现，相同特征对于不同病程患者的重要性有差异。

精神分裂症患者的脑结构随病程改变，有研究者采用 98 例不同病程的精神分裂症患者和 83 例健康对照的皮层厚度数据进行分析，去除了年龄和性别对皮层厚度的影响，使用基于 RBF (Radial Basis Function)核的 SVM 算法进行分类并用留一法验证，获得 81.77%的平均分类正确率，

表 1　文献总结

作者	样本量	模态	算法	特征选择	交叉验证	正确率
Wang et al., 2011	SCZ (32) HC (32)	T1	SVM		十折交叉验证	59.4%~82.8%
Greenstein et al., 2012	SCZ (98) HC (99)	T1	RF			73.7%
Nieuwenhuis et al., 2012	训练集: SCZ (128) HC (111) 测试集: SCZ (155) HC (122)	T1	Linear SVM	结合SVM选特征, 基于以往知识	留一法	训练集: 67.5%~86.8% 测试集: 69.1%~70.6%
Iwabuchi et al., 2013	SCZ (19) HC (20)	T1	Linear SVM		留一法	63.9%~77%
Schnack et al., 2014	训练集: SCZ (66) BD (66) HC (66) 测试集: SCZ (46) BD (47) HC (43)	T1	SVM		Leave-4-out, Leave-2-out	训练集: SCZ vs HC: 90.15% SCZ vs BD: 87.9% HC vs BD: 59.85% 测试集: HC vs SCZ: 76% BD vs SCZ: 66% HC vs BD: 61%
Gould et al., 2014	CD (74) CS (126) HC (134)	T1	Linear SVM		Leave-2-out	HC vs CD: 64%~72% HC vs CS: 59%~67% CD vs CS: 54%~59% HC vs SCZ: 59%~68%
Guo et al., 2016	SCZ (98) HC (83)	T1	SVM-RBF		留一法	81.77%~96.3%
Lu et al., 2016	SCZ (41) HC (42)	T1	SVM	SVM-RFE	留一法	79.5%~88.4%
Salvador et al., 2017	SCZ (128) BD (128) HC (127)	T1	Ridge LR; Lasso LR; Elastic Net Regularization; L0-norm Regularization; SVM; RDA; GPC; RF	单变量 t 值	嵌套交叉验证(外层: 十折交叉验证, 内层: 十折交叉验证)	SCZ vs HC: 58.2%~79.2% SCZ vs BD: 49.2%~66% BD vs HC: 49.1%~65.5%
Chin et al., 2018	训练集: SCZ (84) HC (43) 测试集: SCZ (57) HC (28)	T1	SVM	sequential ROI selection algorithm		训练集: 86.6%~92% 测试集: 83.5%~89.4%
de Pierrefeu et al., 2018	训练集: SCZ (276) HC (330) 测试集: FEP (43) HC (90)	T1	linear SVM, Elastic Net, Enet-TV		二折交叉验证	训练集: 61%~72% 测试集: 61%~73%
Takayanagi et.al., 2011	FEP (52) HC (40)	T1	LDA			81.2%~96.7%
Zanetti, et. al., 2013	FEP (62) HC (62)	T1	SVM	结合 SVM 选特征	留一法	64.3%~73.4%
Dluhoš et al., 2017	FEP (258) HC (222)	T1	SVM		留一法	52%~66%

续表

作者	样本量	模态	算法	特征选择	交叉验证	正确率
Squarcina et al., 2017	FEP (127) HC (127)	T1	SVM-RBF, MKL	结合 SVM 选特征	留一法	68%~85%
Xiao et al., 2017	FEP (163) HC (163)	T1	Linear SVM	两样本 t 检验 (p<0.05)	十折交叉验证	81.8%~85.0%
Winterburn et al., 2017	FEP (50) HC (220) SCZ (179)	T1	LR, SVM (RBF/Linear), LDA		十折交叉验证	训练集： SCZ vs HC：50.0%~71.9% FEP vs HC：50.0%~71.2% 测试集： SCZ vs HC：51.0%~70.8% FEP vs HC：50.6%~73.5%
de Moura et al., 2018	FEP (32) chronic SCZ (143) HC (32)	T1	MLDA		留一法	SCZ vs HC：73.0% FEP vs HC：59.4% sensitivity
Ardekani et al., 2011	训练集：SCZ (25) HC (25) 测试集：SCZ (25) HC (25)	DTI	LDA	两样本 t 检验 (p<0.5)	100 次二折交叉验证	训练集：94%~98% 测试集：96%
Rathi et. al., 2010	FEP (21) HC (20)	DWI	Parzen window classifier, KNN, SVM			69%~86% sensitivity 68%~85% specificity
Mikolas et. al., 2018	FEP (77) HC (77)	DTI	Linear SVM		Leave-2-out	62.34%
Deng et al., 2019	训练集：FEP (52) HC (48) 测试集：FEP (13) HC (12)	DTI	RF	RF-RFE	留一法	训练集：71.0% 测试集：76.0%
Sui et al., 2013	SCZ (35) HC (28)	T1 DTI rs-fMRI	SVM (RBF/Linear), KNN, GNB	mCCA+jICA	十折交叉验证	[a] 单模态：≈(50%~70%) [a] 多模态：≈(60%~80%)
Sui et al., 2014	训练集：SCZ (43) HC (48) 测试集：SCZ (5) HC (5)	T1 rs-fMRI EEG	Linear SVM	mCCA，两样本 t 检验，SVM-RFE	十折交叉验证	训练集： 单模态：66%~93% [a] 多模态：(≥80%)~91% 测试集： 单模态：50%~90% 多模态：80%~100%
Qureshi et al., 2017	SCZ (72) HC (72)	T1 rs-fMRI	SVM (RBF/Linear), ELM, LDA, RF		10 次十折交叉验证	单模态：42.4%~93.0% 多模态：54.9%~99.3%
Tas et al., 2018	SCZ (23) schizo-obsessive (23)	T1 T2	SVM-RBF	ACO	嵌套交叉验证(外层：六折交叉验证，内层：五折交叉验证)	78.26%

续表

作者	样本量	模态	算法	特征选择	交叉验证	正确率
Peruzzo et al., 2015	FEP (23) HC (23)	T1 DTI	SVM, MKL	非参数统计检验(p<0.05)+结合 SVM 选特征	留一法	78.26%~93.48%
Lu et al., 2018	FEP (44) chronic SCZ (44) HC (56)	rs-fMRI T1	SVM	SVM-RFE		[a] 多模态：≥85%
Valli et al., 2016	HR (25) HC (25)	T1 T2	Linear SVM		留一法	54%~72%
Karageorgiou et. al., 2011	SCZ (28) HC (47)	T1 Neuropsychological	stepwise LDA PCA-LDA	Wilk's Λ(p=0.05 保留, p=0.10 删除)	留一法	单模态：70.7%~72% 神经心理和脑影像数据：76%~92%
Chu et al., 2016	SCZ (19) HC (16)	T1 the error rate of the WCST	SVM			82.794%~91.575%
Ebdrup et al., 2018	FEP (46) HC (58)	Cognition data (DART, WAIS III) T1 DTI EEG	naïve Bayes, LR, SVM, decision tree, RF, auto-sklearn	反向特征消除		单模态：49%~56% 认知数据和影像数据：51%~ 68%
Pettersson-Yeo et al., 2013	UHR (19) FEP (19) HC (23)	T1 DTI fMRI Molecular genetics Neuropsychology	SVM		留一法	UHR vs HC：36.84%~68.42% UHR vs FEP：33.33%~76.67% HC vs FEP：44.74%~68.42%
Zarogianni et al., 2017	HR[ill] (17) HR[symp] (17)	T1 behavioural data (RISC score, RAVLT score)	SVM	SVM-RFE	嵌套交叉验证(外层：留一法，内层：留一法)	单模态：88% 行为和影像数据：94.1%

注：SCZ：精神分裂症患者；HC：健康对照；FEP：首发精神分裂症患者；UHR：超高危人群；HR：高危人群，其中，HR[ill]是转变为精神分裂症的高危人群，HR[symp]是未转变为精神分裂症的高危人群；BD：双相情感障碍患者；DTI：弥散张量成像；T1：T1 加权像；DWI：diffusion weighted imaging (弥散加权成像)；fMRI：功能磁共振成像，rs-fMRI 为静息态功能磁共振成像；EEG：脑电波；反向特征消除：backwards elimination feature selection；LDA：Linear Discriminant Analysis (线性判别式分析)；SVM：Support Vector Machine (支持向量机)，其中 SVM-RBF 是核函数为高斯核的支持向量机，Linear SVM 为线性支持向量机；ELM：extreme learning machine (极限学习机)；naïve Bayes：朴素贝叶斯；decision tree：决策树；kNN：k-Nearest Neighbor (k 最近邻)；GNB：Gaussian Naive Bayes (高斯朴素贝叶斯)；MKL：Multiple Kernel Learning (多核学习算法)；Elastic Net Regularization：弹性网络正则化，Enet-TV：ElasticNet-Total Variation；L0-norm Regularization：L0-norm 正则化；RDA：Regularized Discriminant function Analysis (正则化判别函数分析)；GPC：Gaussian Process Classifier (高斯过程分类器)；RF：Random Forests (随机森林)；LR：Logistic Regressions (逻辑回归)，其中，Lasso LR 为 Lasso 回归，Ridge LR 为岭回归；SVM-RFE：Support Vector Machine-Recursive Feature Elimination (基于支持向量机的递归特征消除方法)；DART：Danish Adult Reading Test；WAIS III：Wechsler Adult Intelligence Scale III，韦氏成人智力量表第三版；MLDA：Maximum Uncertainty Linear Discriminant Analysis；PCA：Principal Component Analysis (主成分分析)；ACO：ant colony optimization (蚁群优化)；schizo-obsessive：伴随强迫症状的精神分裂症；cognitive deficit subtype：精神分裂症的认知缺陷亚型；cognitively spared subtype：精神分裂症的认知正常亚型

a：文章中没有明确给出正确率

77.55%的平均敏感性和 86.75%的平均特异性，进一步研究发现病程小于两年的患者与健康对照的分类正确率最高(96.3%，特异性为 98.8%、敏感性为 88%)；患者尤其是长病程患者的颞叶−边缘系统和额叶−顶叶的皮层厚度减少，但枕叶皮层厚度增加(Guo et al., 2016)。

在分类过程中，有研究者提出与使用全脑灰质密度数据相比，利用特征选择方法选取部分脑区的灰质密度作为特征能提高分类正确率。Chin 等(2018)使用序列感兴趣区选择方法(sequential region-of-interest (ROI) selection)在 64 个脑区的灰质密度数据中选择最优特征，结果发现使用该特征选择方法可将 84 个精神分裂症患者和 43 个健康对照(训练集)的分类正确率从 86.6%提高到 92.0%，将 57 个精神分裂症患者和 28 个健康对照(测试集)的分类正确率从 83.5%提高到 89.4%，同时筛选出包含 7 个 ROI 的最佳分类子集，包括梭状回、额中回、额下回、颞上回、额上回、左丘脑和左侧脑室。还有研究者比较不同特征选择方法对分类正确率的影响。Nieuwenhuis 等人(2012)采用基于体素的形态学分析(Voxel-Based Morphometry, VBM)提取脑结构像每个体素的灰质密度数据，用线性回归去除年龄、性别和利手的影响，将残差作为原始输入数据，分别采用嵌入法和基于先验知识选取特征，前者使用 SVM 算法训练模型，选取权重绝对值前 10%的特征作为输入特征，后者基于以往研究去除受药物影响的纹状体的灰质密度，将余下脑区的灰质密度作为输入特征。使用线性 SVM 算法在 239 个被试(128 个患者和 111 个健康对照)中建立模型，使用留一法交叉验证。结果发现采用嵌入法选取特征建立的模型具有较好的分类表现：71.4%的平均正确率，73.4%的平均敏感性和 69.4%的平均特异性；最后在 277 个被试(155 个患者和 122 个健康对照)的数据中进行验证，得到 70.4%的分类正确率，67.1%的敏感性和 73.8%的特异性；研究发现患者额叶、颞上回和海马的灰质密度降低，而基底神经节和左侧枕叶灰质密度增加。

除了使用不同的特征选择方法，研究者还探究不同算法对分类正确率的影响，Salvador 等人(2017)使用 127 个健康对照和 128 个精神分裂症患者的灰质体积探究不同算法的分类性能。该研究采用的算法包括 Ridge，Lasso，弹性网络(Elastic Net)，L0 范数正则化逻辑回归(L0 Norm Regularized Logistic Regressions)，线性 SVM，正则化判别分析(Regularized Discriminant Analysis)，随机森林(Random Forests)和高斯过程分类(Gaussian Process Classifier)，发现使用这 8 种算法得到的分类结果并无较大差异，分类正确率在 74.1%到 76%之间。

现有研究不仅将机器学习应用于区分精神分裂症患者和健康对照，而且还尝试区分不同亚型精神分裂症患者。有研究者使用机器学习区分有无强迫症状的精神分裂症患者，在 23 例无强迫症状的精神分裂症患者和 23 例有强迫症状的精神分裂症患者的样本中发现中央旁小叶(包括辅助运动区)、中扣带回(middle cingulate gyrus)和额叶−皮层下(fronto-subcortical)区域的灰质密度能很好地区分两组被试，正确率达到 78.26% (Tas et al., 2018)。

2.1.2 首发精神分裂症

因受到病程和长期服用药物的影响，慢性精神分裂症患者的脑结构可能发生一定的改变，影响机器学习研究的分类表现。首发精神分裂症患者较少受到药物和病程的干扰，对这一群体的研究有利于探讨精神分裂症的神经病理机制。在首发精神分裂症患者的分类研究中，总样本量最小仅为 42，最大达到 480；分类正确率在 44.74%与 96.7%之间。在特征方面，额叶(如额上回、额中回和额极等)、颞中回、梭状回和舌回的脑结构特征能提高分类正确率。

分类问题是研究者首要关注的问题，在 127 个首发患者和 127 个健康对照的分类研究中，Squarcina 等人(2017)提取 68 个感兴趣区的皮层厚度数据，分别对每个感兴趣区应用 SVM 算法，根据分类结果选取 5 个最优特征；利用最优特征建立模型，使用 SVM 或多核学习(Multiple Kernel Learning，MKL)算法进行训练和预测，在额上回，额中回喙部和颞中回取得 85%的最高正确率。

在首发精神分裂症患者中有研究者探究了不同算法应用于灰质密度和皮层厚度数据的分类表现。Winterburn 等人(2017)使用逻辑回归(Logistic Regression，LR)、SVM (包括线性 SVM 和 SVM-RBF)和线性判别分析(Linear Discriminant Analysis，LDA)三种算法进行分类，脑结构像数据包括灰质密度和皮层厚度，数据样本包括 50 例首发精神分裂症患者和 50 例健康对照，研究者首先将样本分

为 2/3 的训练集和 1/3 的测试集, 应用线性 SVM 算法于灰质密度数据时, 分类正确率最高可达 69.6% (敏感性 77.8%, 特异性 57.9%); 使用 SVM-RBF 算法和皮层厚度数据进行分类, 最高分类正确率可达 73.5% (敏感性 58.8%, 特异性 88.2%)。

2.1.3　高危人群

与首发精神分裂症患者相比, 识别高危人群与探究其脑结构有利于解释精神分裂症的发生机制、预测精神疾病的发生。目前利用脑结构像对高危人群的分类研究较少, 总样本量均不超过 50, 分类正确率在 36.84%与 72%之间。其中, 有研究者使用灰质体积数据和 SVM 算法对 25 名高危个体和 25 名健康对照个体进行分类, 应用留一法进行交叉验证, 获得 72%的分类正确率(敏感性 68%, 特异性 76%), 双侧海马、海马旁回、壳核、额上回、额中回、梭状回和顶下小叶, 左侧颞下回、右侧颞上回、左侧楔前叶和左侧小脑等区域具有较高的特征权重(Valli et al., 2016)。

2.2　基于弥散张量成像的单模态机器学习研究

弥散张量成像是另一种常用的脑结构成像。水分子在脑中受到白质纤维(如髓鞘和轴突膜等)的约束, 各个方向上的弥散不同, 呈现各向异性。弥散张量成像利用水分子弥散的各向异性成像, 用于探究脑白质纤维束的位置和方向(Meoded, Poretti, Mori, & Zhang, 2017)。现有基于 DTI 的机器学习分类研究较少, 样本量在 42 到 154 之间, 研究常采用的特征包括各向异性分数(Fractional Anisotropy, FA)、轴向弥散张量(Axial Diffusivity, AD)、垂直弥散张量(Radial Diffusivity, RD)、平均弥散率(Mean Diffusivity, MD)等, 从机器学习的分类表现上看, 现有研究的分类正确率最低为 60%, 最高可达 98%。

有研究者将 50 名精神分裂症患者和 50 名健康对照分为训练集和测试集, 基于 DTI 数据计算 FA 和 MD 值, 使用 PCA 降维并在数据中进行独立样本 t 检验, 移除 P 值大于 0.5 的特征, 利用 Fisher 线性判别分析训练模型。结果发现使用 FA 值构建的模型在测试集中获得 94%正确率(敏感性 96%, 特异性 92%), 有区分能力的特征集中在大脑深部白质区域, 如双侧外囊(external capsule); 使用 MD 值构建的模型在测试集中的正确率为 98% (敏感性 96%, 特异性 100%), 影响正确率的特征集中在皮层和脑室; 结合 FA 和 MD 数据进行分类并没有显著改变模型的分类正确率(正确率、敏感性和特异性均为 96%) (Ardekani et al., 2011)。

另外, 有研究者基于 DTI 数据计算 18 条主要白质纤维束的 FA、MD、RD 和 AD 值, 使用基于随机森林的递归特征消除法(Random Forest-Recursive Feature Elimination, RF-RFE)选择特征, 采用随机森林算法在 52 个首发患者和 48 个健康对照的数据中进行训练和测试, 得到 71%的分类正确率(敏感性 67.3%, 特异性 75%); 用最终模型区分 13 个首发患者和 12 个健康对照得到 76%的分类正确率(敏感性 76.9%, 特异性 75%), 具有区分能力的特征位于连合纤维、小脑−丘脑−皮层回路(the cerebello-thalamo-cortical circuits)和长纤维(Deng et al., 2019)。

Pettersson-Yeo 等人(2013)利用 FA 值分别对 19 个高危个体和 23 个健康对照、19 个首发患者和 23 个健康对照进行分类, 使用 SVM 算法均得到 65.79%的分类正确率, 其敏感性均为 68.42%, 特异性均为 63.16%。

2.3　基于多模态脑影像数据的机器学习研究

除了单模态数据, 现有研究还尝试结合多种模态脑影像数据对精神分裂症患者或高危人群与健康对照进行分类。目前多模态脑影像机器学习分类研究有 8 篇, 总样本量在 46 到 144 之间, 分类正确率在 54.9%到 100%之间, 常与 T1 加权像数据结合的脑影像数据有静息态功能磁共振成像(resting state functional Magnetic Resonance Imaging, rs-fMRI)数据和 DTI 数据。

Peruzzo 等人(2015)结合 T1 加权像和 DTI 数据, 采用非参数统计检验和 SVM 算法选取特征, 采用多核学习 MKL 算法对 23 个首发精神分裂症患者和 23 个健康对照进行分类, 正确率最高达到 93.5%, 影响正确率的特征包括杏仁核、额上回、额下回、苍白球、海马旁回的灰质体积和侧脑室体积, 额上回、额中回、额下回、梭状回、海马旁回、颞中回、眶额皮层和脑岛的皮层厚度, 钩束、扣带、小脑下角和胼胝体压部的弥散张量值。此外, Sui 等人(2013)提出一种多模态融合的方法——mCCA (multi-set Canonical Correlation Analysis)和 jICA (joint Independent Component Analysis), 将 rs-fMRI、结构像和 DTI 的数据两两结合, 对 35 名精神分裂症患者和 28 名健康对照个体进行

分类，发现单模态数据的分类正确率在 50%~70%之间，多模态数据正确率在 60%~80%之间。另外，该研究团队还基于 rs-fMRI、T1 加权像和脑电数据(Electroencephalogram, EEG)采用独立样本 t 检验，mCCA 和 SVM-RFE 选择特征，应用 SVM 算法和十折交叉验证方法在 43 个精神分裂症患者和 48 个健康对照中得到 91%的分类正确率，并且以 100%的正确率成功区分 5 个精神分裂症患者和 5 个健康对照个体(Sui et al., 2014)。

虽然现阶段大部分多模态脑影像的机器学习研究分类正确率较高，但此类研究的样本量通常较小。因此，多模态脑影像数据对分类结果的影响仍需要进一步探讨。

2.4 脑影像数据和行为数据相结合的机器学习研究

除结合不同模态的脑影像数据以外，目前也有将脑影像数据和行为数据相结合作为分类特征的机器学习研究，现有研究建议将脑影像数据和行为数据结合能提高分类正确率。

有研究将灰质、白质、脑脊液体积与威斯康星卡片分类测试(Wisconsin card sorting test, WCST)的错误率结合，利用独立成分分析处理数据，使用 SVM 算法在 19 个精神分裂症患者和 16 个健康对照的数据中取得 91.58%的分类正确率，证明 WCST 的错误率和脑影像数据结合能有效区分患者和健康对照(Chu et al., 2016)。Zarogianni 等人(2017)预测 34 名具有家族遗传史的高危个体是否会转变为精神分裂症患者，应用 SVM-RFE 算法在各脑区的灰质密度数据、Rust Inventory of Schizotypal Cognitions (RISC)问卷分数和雷氏听觉语言学习测验(Rey Auditory Verbal Learning Test, RAVLT)得分中进行特征选择，用 SVM 算法进行预测，分类正确率达到 94%，高于仅使用灰质密度数据的分类正确率(88%)。Ebdrup 等人(2018)对 46 个首发精神分裂症患者和 58 个健康对照进行分类，结合行为数据(包括韦氏智力量表第三版和丹麦成人阅读测试(Danish Adult Reading Test))和 T1 加权像、DTI 和 EEG 的数据仅获得 51%~68%的分类正确率，低于仅使用认知数据的分类结果(60%~69%)，稍高于单独使用单模态数据的分类结果(49%~56%)。因此，现有研究提示结合行为数据和脑影像数据的分类结果可能优于单模态数据的分类结果，但结合这两类数据能否取得高正确率仍有待进一步探讨。

3 机器学习研究的局限

将机器学习应用于精神分裂症领域，利用脑结构特征对精神分裂症患者、高危人群与健康对照进行分类，有助于理解疾病的生理病理机制。研究数量逐年递增，但仍具有一定局限，特别是在研究样本量和模型泛化能力上仍存在问题，亟待解决。

3.1 样本量

图 2 总结了现有精神分裂症和高危人群中机

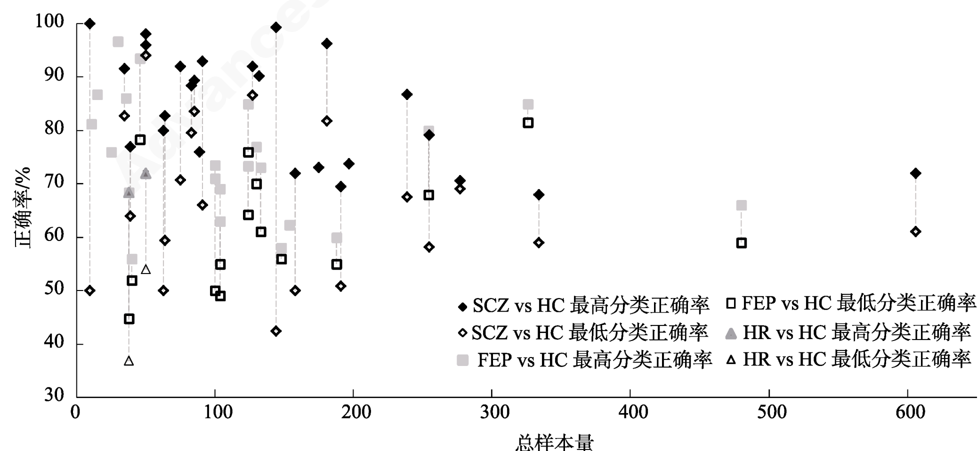


图 2 精神分裂症患者、高危人群基于脑结构像分类研究的总样本量和分类正确率

注：SCZ：精神分裂症; HC：健康对照; FEP：首发精神分裂症; HR：高危人群。实心点代表最高分类正确率; 空心点代表最低分类正确率; 虚线连接同一研究中相同样本量的最高与最低分类正确率。

器学习研究采用的总样本量以及分类的最高和最低正确率。如图 2 所示，样本量在 50 到 150 时，最高分类正确率较高，样本量大于 400 时，最高正确率只有 60%~72%。出现小样本分类正确率高而大样本分类正确率低的情况可能有如下原因：1)样本量小导致模型不稳定和测试不准确。训练样本量过小容易导致训练不充分，造成模型过拟合或不稳定；测试样本量过小，容易出现过高或过低的分类正确率，即使在小样本测试中得到高正确率仍难以验证一个模型的性能。有研究表明，验证一个模型的有效性至少需要 75 到 100 个测试样本(Beleites, Neugebauer, Bocklitz, Krafft, & Popp, 2013)；2)现有大样本研究的分类特征远大于样本量，研究中未进行特征选择，容易造成模型过拟合，降低分类正确率; 3)大样本研究数据来自多个中心，患者的异质性可能更高，从而降低分类正确率。

样本量不足是现阶段精神分裂症领域机器学习研究面临的一大问题(Madsen et al., 2018)。样本量不足的主要原因包括磁共振扫描成本高、精神分裂症的特殊性和数据共享困难。高昂的磁共振成像扫描费用和较长的扫描时间导致研究者难以在短时间内收集大量脑影像数据。精神分裂症患者受到临床症状和药物副作用的影响，在采集脑影像数据时需要耗费更多的人力和物力，采集难度大，因此与精神分裂症有关的脑成像研究样本量通常较小。除此之外，多中心数据共享困难是也是可能的原因之一，各中心扫描仪器的型号和参数各不相同，并且存在数据管理、伦理和隐私保护的问题。

未来多个研究机构/中心可以通过制定详细协议保障被试隐私和数据安全，统一扫描参数和入组标准，在不违背伦理的情况下实现多中心数据共享，增大样本量。研究者在纳入多中心数据时，需要保证入组标准尽可能一致；在处理大样本数据时，进行数据清洗，利用降维和特征选择的方法保留有用特征，保证数据质量，得到更加可信的结果，从而利用大样本数据考察脑结构特征的有效性。有研究者提出可以通过合并多中心模型解决样本量问题。研究者分别在 4 个中心收集首发精神分裂症患者和健康对照的脑结构数据，分别使用 4 个中心的数据训练模型，利用这 4 个模型生成最终的元模型(the meta-model)。研究发现该元模型与基于所有数据生成的模型有较高的相似性(Dluhoš et al., 2017)。因此，合并多中心模型为解决样本无法共享提供了一种新思路。

3.2　泛化能力

机器学习的目标是使训练得到的模型能够很好地适用于新样本，即具有较好的泛化能力。除了样本量不足会影响模型泛化能力外，机器学习训练过程中出现信息泄露也会影响泛化能力。常见的信息泄露的情况是使用同一批样本调整参数和测试结果，在特征选择时提前使用测试集的标签，如在所有样本中进行独立样本 t 检验选取显著脑区，再进行训练集、测试集和验证集的拆分。提前使用测试集的信息，虽然能得到高正确率，但是不能反映模型真正的泛化能力。研究者在今后研究中应避免信息泄露并使用嵌套交叉验证(Madsen et al., 2018; Dwyer et al., 2018)，使评估结果更加精确，提高模型泛化能力。为测试泛化能力，Dwyer 等人(2018)建议未来可设置第三方检测机构，收集各中心或机构的数据，一旦形成新模型，研究者可移交第三方机构，利用其数据进行测试，从而保证较好的泛化能力。

4　未来研究展望

未来的精神分裂症脑影像机器学习研究可以考虑在特征中加入用药情况、临床症状和社会人口学等信息。Kambeitz 等人(2015)在精神分裂症患者和健康对照的脑影像机器学习研究中，将特异性和敏感性作为效应量，考察年龄、阳性症状与阴性症状比率、抗精神病药物剂量对分类结果的影响。该元分析发现，年龄越大机器学习分类的敏感性越高；阳性症状与阴性症状比率越大特异性越高；抗精神病药物剂量越大特异性越高。因此，在区分不同病程的患者或探究预后情况时，可以纳入临床信息和社会人口学信息，提高分类表现。

近年来，随着计算机技术的高速发展，深度学习开始重新获得人们关注。深度学习是以人工神经网络为架构，对数据进行表征学习的算法，在图像、视频、语音等多个领域均表现优异(Bengio, Courville, & Vincent, 2013; Lecun, Bengio, & Hinton, 2015; Schmidhuber, 2015)。Pinaya 等人(2016)使用深度信念网络(Deep Belief Network)，在 143 例精神分裂症患者和 83 例健康对照中得到了 73.6%的

分类正确率(敏感性 76.37%，特异性 70.74%)，稍高于使用 SVM 算法的正确率(68.1%)。Dwyer 等人(2018)指出，深度学习需要大样本数据来避免维度灾难和过拟合。因此，研究者可以在获得大样本数据的基础上，尝试使用深度学习的方法进行分类，探究精神分裂症的内在机制。

现有研究的数据处理流程和结果报告都不相同，由于缺乏具体信息和使用方法的差异，不同研究之间很难进行比较。据此 Riley (2019)指出将机器学习应用于不同学科领域中时，应建立统一的数据分析和结果报告标准。

5 总结

机器学习技术从数据中提取有用信息，可以协助研究者在个体水平上进行预测，探究精神分裂症的特点和内在机制，帮助寻找生物标记物、实现精神疾病的早期识别。本文主要从脑结构像出发，综述了单模态、多模态脑影像和结合行为数据与脑影像的机器学习研究。单模态脑影像的机器学习分类正确率在 36.84%到 98%之间，多基于 T1 加权像，结果提示额叶和颞叶的脑结构特征可以有效区分精神分裂症患者或高危人群和健康对照；基于 DTI 的机器学习研究多采用弥散张量值为特征，但因其样本量小且现有研究少，各研究者提出的能提高分类正确率的脑区或纤维特征大不相同，尚未达成统一意见。多模态脑影像的分类正确率在 54.9%~100%之间，多数研究的分类正确率高，但样本量小，多模态数据是否显著优于单模态数据并有效提高分类正确率，有待研究者进一步探讨。行为数据和脑影像数据结合的分类效果优于单模态数据的分类效果，但能否显著提高分类正确率仍待进一步验证。目前精神分裂症领域的机器学习研究具有样本量不足和泛化能力弱等局限，未来还需要在多中心的大样本中进行模型训练和验证，以求获得更可靠的模型和更有效的特征，才能在临床实践中发挥更大作用。

参考文献

周志华. (2016). *机器学习*. 清华大学出版社.

Ardekani, B. A., Tabesh, A., Sevy, S., Robinson, D. G., Bilder, R. M., & Szeszko, P. R. (2011). Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. *Human Brain Mapping*, *32*(1), 1–9. https://doi.org/10.1002/hbm.20995

Bakhshi, K., & Chance, S. A. (2015). The neuropathology of schizophrenia: A selective review of past studies and emerging themes in brain structure and cytoarchitecture. *Neuroscience*. *33*, 82–102. https://doi.org/10.1016/j.neuroscience.2015.06.028

Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C., & Popp, J. (2013). Sample size planning for classification models. *Analytica Chimica Acta*, *760*, 25–33. doi:10.1016/j.aca.2012.11.007

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Borgwardt, S., Koutsouleris, N., Aston, J., Studerus, E., Smieskova, R., Riecher-Rössler, A., & Meisenzahl, E. M. (2013). Distinguishing prodromal from first-episode psychosis using neuroanatomical single-subject pattern recognition. *Schizophrenia Bulletin*, *39*(5), 1105–1114. https://doi.org/10.1093/s chbul/sbs095

Carpenter, W. T., & Buchanan, R. W. (1994). Schizophrenia. *New England Journal of Medicine*, *330*(10), 681–690. https://doi.org/10.1056/NEJM199403103301006

Chin, R., You, A. X. B., Meng, F. W., Zhou, J., & Sim, K. (2018). Recognition of schizophrenia with regularized support vector machine and sequential region of interest selection using structural magnetic resonance imaging. *Scientific Reports*, *8*(1), 13858. https://doi.org/10.1038/s41598-018-32290-9

Chu, W.-L., Huang, M.-W., Jian, B.-L., Hsu, C.-Y., & Cheng, K.-S. (2016). A correlative classification study of schizophrenic patients with results of clinical evaluation and structural magnetic resonance images. *Behavioural Neurology*, *2016*, 1–11. https://doi.org/10.1155/2016/7849526

de Moura, A. M., Pinaya, W. H. L., Gadelha, A., Zugman, A., Noto, C., Cordeiro, Q., … Sato, J. R. (2018). Investigating brain structural patterns in first episode psychosis and schizophrenia using MRI and a machine learning approach. *Psychiatry Research - Neuroimaging*, *275*(March), 14–20. https://doi.org/10.1016/j.pscychresns.2018.03.003

Deng, Y., Hung, K. S. Y., Lui, S. S. Y., Chui, W. W. H., Lee, J. C. W., Wang, Y., … Cheung, E. F. C. (2019). Tractography-based classification in distinguishing patients with first-episode schizophrenia from healthy individuals. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *88*, 66–73. https://doi.org/10.1016/j.pnpbp.2018.06.010

de Pierrefeu, A., Löfstedt, T., Laidi, C., Hadj-Selem, F., Bourgin, J., Hajek, T., … Duchesnay, E. (2018). Identifying a neuroanatomical signature of schizophrenia, reproducible across sites and stages, using machine learning with

structured sparsity. *Acta Psychiatrica Scandinavica*. *138*(6), 571–580. https://doi.org/10.1111/acps.12964

Dluhoš, P., Schwarz, D., Cahn, W., van Haren, N., Kahn, R., Španiel, F., … Schnack, H. (2017). Multi-center machine learning in imaging psychiatry: A meta-model approach. *NeuroImage*, *155*(April), 10–24. https://doi.org/10.1016/j.neuroimage.2017.03.027

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, *14*(1), 91–118. https://doi.org/10.1146/annurev-clinpsy-032816-045037

Dyrba, M., Grothe, M., Kirste, T., & Teipel, S. J. (2015). Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. *Human Brain Mapping*, *36*(6), 2118–2131. https://doi.org/10.1002/hbm.22759

Ebdrup, B. H., Axelsen, M. C., Bak, N., Fagerlund, B., Oranje, B., Raghava, J. M., … Glenthøj, B. Y. (2018). Accuracy of diagnostic classification algorithms using cognitive-, electrophysiological-, and neuroanatomical data in antipsychotic-naïve schizophrenia patients. *Psychological Medicine*. https://doi.org/10.1017/S0033291718003781

Gould, I. C., Shepherd, A. M., Laurens, K. R., Cairns, M. J., Carr, V. J., & Green, M. J. (2014). Multivariate neuroanatomical classification of cognitive subtypes in schizophrenia: A support vector machine learning approach. *NeuroImage: Clinical, 6*, 229-236.

Greenstein, D., Malley, J. D., Weisinger, B., Clasen, L., & Gogtay, N. (2012). Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls. *Frontiers in Psychiatry*, *3*(June), 1–12. https://doi.org/10.3389/fpsyt.2012.00053

Guo, S., Palaniyappan, L., Liddle, P. F., & Feng, J. (2016). Dynamic cerebral reorganization in the pathophysiology of schizophrenia: A MRI-derived cortical thickness study. *Psychological Medicine*, *46*(10), 2201–2214. https://doi.org/10.1017/S0033291716000994

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning, 46*(1-3), 389–422. https://doi.org/10.1023/A:1012487302797

Iwabuchi, S. J., Liddle, P. F., & Palaniyappan, L. (2013). Clinical utility of machine-learning approaches in schizophrenia: Improving diagnostic confidence for translational neuroimaging. *Frontiers in Psychiatry*, *4*(August), 1–9. https://doi.org/10.3389/fpsyt.2013.00095

Kambeitz, J., Kambeitz-Ilankovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., … Koutsouleris, N. (2015). Detecting neuroimaging biomarkers for schizophrenia: A meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology, 40*(7), 1742–1751. https://doi.org/10.1038/npp.2015.22

Karageorgiou, E., Schulz, S. C., Gollub, R. L., Andreasen, N. C., Ho, B. C., Lauriello, J., ... & Georgopoulos, A. P. (2011). Neuropsychological testing and structural magnetic resonance imaging as diagnostic biomarkers early in the course of schizophrenia and related psychoses. *Neuroinformatics, 9*(4), 321–333.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521,* 436–444. https://doi.org/10.1038/nature14539

Lu, X. B., Yang, Y. Z., Wu, F. C., Gao, M. J., Xu, Y., Zhang, Y., … Wu, K. (2016). Discriminative analysis of schizophrenia using support vector machine and recursive feature elimination on structural MRI images. *Medicine (United States), 95*(30), e3973. https://doi.org/10.1097/MD.0000000000003973

Lu, X.-B., Zhang, Y., Yang, D.-Y., Yang, Y.-Z., Wu, F.-C., Ning, Y.-P., & Wu, K. (2018). Analysis of first-episode and chronic schizophrenia using multi-modal magnetic resonance imaging. *European Review for Medical and Pharmacological Sciences*, *22*(19), 6422–6435. https://doi.org/10.26355/eurrev_201810_16055

Madsen, K. H., Krohne, L. G., Cai, X.-L., Wang, Y., & Chan, R. C. K. (2018). Perspectives on machine learning for classification of Schizotypy using fMRI data. *Schizophrenia bulletin, 44*(suppl_2), S480-S490. https://doi.org/10.1093/schbul/sby026

Meoded, A., Poretti, A., Mori, S., & Zhang, J. (2017). Diffusion Tensor Imaging (DTI). *Reference Module in Neuroscience and Biobehavioral Psychology*, (March 2016), 1–11. https://doi.org/10.1016/B978-0-12-809324-5.02472-X

Mikolas, P., Hlinka, J., Skoch, A., Pitra, Z., Frodl, T., Spaniel, F., & Hajek, T. (2018). Machine learning classification of first-episode schizophrenia spectrum disorders and controls using whole brain white matter fractional anisotropy. *BMC Psychiatry*, *18*(1), 97. https://doi.org/10.1186/s12888-018-1678-y

Mitchell, T. (1997). *Machine Learning*. McGraw-Hill, New York.

Nieuwenhuis, M., van Haren, N. E. M., Hulshoff Pol, H. E., Cahn, W., Kahn, R. S., & Schnack, H. G. (2012). Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *NeuroImage*, *61*(3), 606–612. https://doi.org/10.1016/j.neuroimage.2012.03.079

Palaniyappan, L. (2017). Progressive cortical reorganisation: A framework for investigating structural changes in schizophrenia. *Neuroscience and Biobehavioral Reviews, 79,* 1–3. https://doi.org/10.1016/j.neubiorev.2017.04.028

Peng, B., Lu, J., Saxena, A., Zhou, Z. Y., Zhang, T., Wang, S.

H., & Dai, Y. K. (2017). Examining brain morphometry associated with self-esteem in young adults using multilevel-ROI-features-based classification method. *Frontiers in Computational Neuroscience*, *11*(May), 1–10. https://doi.org/10.3389/fncom.2017.00037

Peruzzo, D., Castellani, U., Perlini, C., Bellani, M., Marinelli, V., Rambaldelli, G., … Brambilla, P. (2015). Classification of first-episode psychosis: A multi-modal multi-feature approach integrating structural and diffusion imaging. *Journal of Neural Transmission, 122*(6), 897–905. https://doi.org/10.1007/s00702-014-1324-x

Pettersson-Yeo, W., Benetti, S., Marquand, A. F., Dell'Acqua, F., Williams, S. C. R., Allen, P., … Mechelli, A. (2013). Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. *Psychological Medicine*, *43*(12), 2547–2562. https://doi.org/10.1017/S003329171300024X

Pinaya, W. H. L., Gadelha, A., Doyle, O. M., Noto, C., Zugman, A., Cordeiro, Q., … Sato, J. R. (2016). Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. *Scientific Reports*, *6*(1), 38897. https://doi.org/10.1038/srep38897

Qureshi, M. N. I., Oh, J., Cho, D., Jo, H. J., & Lee, B. (2017). Multimodal discrimination of schizophrenia using hybrid weighted feature concatenation of brain functional connectivity and anatomical features with an extreme learning machine. *Frontiers in Neuroinformatics*, *11*(September), 1–14. https://doi.org/10.3389/fninf.2017.00059

Rathi, Y., Malcolm, J., Michailovich, O., Goldstein, J., Seidman, L., McCarley, R. W., & Westin, C.-F. (2010). Biomarkers for identifying first-episode schizophrenia patients using diffusion weighted imaging. *Med Image Comput Comput-Assist Interv, 13,* 657–665.

Riley, P. (2019). Three pitfalls to avoid in machine learning. *Nature, 572*(7767), 27–29.

Salvador, R., Radua, J., Canales-Rodríguez, E. J., Solanes, A., Sarró, S., Goikolea, J. M., … Pomarol-Clote, E. (2017). Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS ONE*, *12*(4), e0175683. https://doi.org/10.1371/journal.pone.0175683

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development, 3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61,* 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Schnack, H. G., Nieuwenhuis, M., van Haren, N. E. M., Abramovic, L., Scheewe, T. W., Brouwer, R. M., … Kahn, R. S. (2014). Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *NeuroImage*, *84*, 299–306.

Squarcina, L., Castellani, U., Bellani, M., Perlini, C., Lasalvia, A., Dusi, N., … Brambilla, P. (2017). Classification of first-episode psychosis in a large cohort of patients using support vector machine and multiple kernel learning techniques. *NeuroImage*, *145*(Pt B), 238–245. https://doi.org/10.1016/j.neuroimage.2015.12.007

Sui, J., Castro, E., He, H., Bridwell, D., Du, Y. H., Pearlson, G. D., … Calhoun, V. D. (2014). Combination of fMRI-sMRI-EEG data improves discrimination of schizophrenia patients by ensemble feature selection. *Conference Proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2014*(2), 3889–3892. https://doi.org/10.1109/EMBC.2014.6944473

Sui, J., He, H., Yu, Q. B., Chen, J. Y., Rogers, J., Pearlson, G. D., … Calhoun, V. D. (2013). Combination of resting state fMRI, DTI, and sMRI data to discriminate schizophrenia by N-way MCCA + jICA. *Frontiers in Human Neuroscience*, *7*(May), 1–14. https://doi.org/10.3389/ fnhum.2013.00235

Takayanagi, Y., Takahashi, T., Orikabe, L., Mozue, Y., Kawasaki, Y., Nakamura, K., ... Suzuki, M. (2011). Classification of first-episode schizophrenia patients and healthy subjects by automated MRI measures of regional brain volume and cortical thickness. *PloS ONE, 6*(6), e21047.

Tas, C., Mogulkoc, H., Eryilmaz, G., Gogcegoz-Gul, I., Erguzel, T. T., Metin, B., & Tarhan, N. K. (2018). Discriminating schizophrenia and schizo-obsessive disorder: A structural MRI study combining VBM and machine learning methods. *Neural Computing and Applications*, *29*(2), 377–387. https://doi.org/10.1007/s00521-016-2451-0

Valli, I., Marquand, A. F., Mechelli, A., Raffin, M., Allen, P., Seal, M. L., & McGuire, P. (2016). Identifying individuals at high risk of psychosis: Predictive utility of support vector machine using structural and functional MRI data. *Frontiers in Psychiatry*, *7*, 52. https://doi.org/10.3389/fpsyt.2016.00052

Wang, L. B., Shen, H., Li, B. J., & Hu, D. W. (2011). Classification of schizophrenic patients and healthy controls using multiple spatially independent components of structural MRI data. *Frontiers of Electrical and Electronic Engineering in China, 6*(2), 353–362. https://doi.org/10.1007/s11460-011-0142-2

Winterburn, J. L., Voineskos, A. N., Devenyi, G. A., Plitman, E., de la Fuente-Sandoval, C., Bhagwat, N., … Chakravarty, M. M. (2017). Can we accurately classify schizophrenia patients from healthy controls using magnetic resonance

imaging and machine learning? A multi-method and multi-dataset study. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2017.11.038

Xiao, Y., Yan, Z. H., Zhao, Y. J., Tao, B., Sun, H. Q., Li, F., … Lui, S. (2017). Support vector machine-based classification of first episode drug-naïve schizophrenia patients and healthy controls using structural MRI. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2017.11.037

Zarogianni, E., Storkey, A. J., Johnstone, E. C., Owens, D. G. C., & Lawrie, S. M. (2017). Improved individualized prediction of schizophrenia in subjects at familial high risk, based on neuroanatomical data, schizotypal and neurocognitive features. *Schizophrenia Research*, *181*, 6–12. https://doi.org/10.1016/j.schres.2016.08.027

Zanetti, M. V., Schaufelberger, M. S., Doshi, J., Ou, Y. M., Ferreira, L. K., Menezes, P. R., ... Busatto, G. F. (2013). Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 43*, 116–125.

The classification of schizophrenia based on brain structural features: A machine learning approach

ZHENG Hong[1,2]; PU Cheng-cheng[3]; WANG Yi[1,2]; Raymond C. K. CHAN[1,2]

([1] *Neuropsychology and Applied Cognitive Neuroscience Laboratory, CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing* 100101, *China*)

([2] *Department of Psychology, University of Chinese Academy of Sciences, Beijing* 100049, *China*)

([3] *Peking University Sixth Hospital, Peking University Institute of Mental Health, NHC Key Laboratory of Mental Health (Peking University), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Beijing* 100191, *China*)

Abstract: Machine learning is a promising approach for mental disorders. In recent years, machine learning based on T1 weighted imaging and Diffusion Tensor Imaging (DTI) data has been used to investigate the psychopathology and underlying mechanisms of schizophrenia patients and high-risk population. The findings from the previous literature suggest that structural features of frontal lobe and temporal lobe can improve classification performance. In addition, the combination of behavioural performances and the features of brain structure is superior to the single-modality structural images on classification accuracy. However, the existing empirical studies classifying schizophrenia patients or high-risk population from controls are limited in sample size and generalization ability.

Key words: structural Magnetic Resonance Imaging; Diffusion Tensor Imaging; machine learning; schizophrenia; high risk population